

EXHIBIT 5

Enzymatic Amplification of β -Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia

Randall K. Saiki, Stephen Scharf, Fred Faloona, Kary B. Mullis
Glenn T. Horn, Henry A. Erlich, Norman Arnheim

Recent advances in recombinant DNA technology have made possible the molecular analysis and prenatal diagnosis of several human genetic diseases. Fetal DNA obtained by amniocentesis or chorionic villus sampling can be analyzed by restriction enzyme digestion, with subsequent electrophoresis, Southern transfer, and specific hybridization to cloned gene or oligonucleotide probes. With

This disease results from homozygosity of the sickle-cell allele (β^S) at the β -globin gene locus. The S allele differs from the wild-type allele (β^A) by substitution of an A in the wild-type to a T at the second position of the sixth codon of the β chain gene, resulting in the replacement of a glutamic acid by a valine in the expressed protein. For the prenatal diagnosis of sickle cell anemia, DNA ob-

Abstract. Two new methods were used to establish a rapid and highly sensitive prenatal diagnostic test for sickle cell anemia. The first involves the primer-mediated enzymatic amplification of specific β -globin target sequences in genomic DNA, resulting in the exponential increase (220,000 times) of target DNA copies. In the second technique, the presence of the β^A and β^S alleles is determined by restriction endonuclease digestion of an end-labeled oligonucleotide probe hybridized in solution to the amplified β -globin sequences. The β -globin genotype can be determined in less than 1 day on samples containing significantly less than 1 microgram of genomic DNA.

polymorphic DNA markers linked genetically to a specific disease locus, segregation analysis must be carried out with restriction fragment length polymorphisms (RFLP's) found to be informative by examining DNA from family members (1, 2).

Many of the hemoglobinopathies, however, can be detected by more direct methods in which analysis of the fetus alone is sufficient for diagnosis. For example, the diagnosis of hydrops fetalis (homozygous α -thalassemia) can be made by documenting the absence of any α -globin genes by hybridization with an α -globin probe (3-5). Homozygosity for certain β -thalassemia alleles can be determined in Southern transfer experiments by using oligonucleotide probes that form stable duplexes with the normal β -globin gene sequence but form unstable hybrids with specific mutants (6, 7).

Sickle cell anemia can also be diagnosed by direct analysis of fetal DNA.

tained by amniocentesis or chorionic villus sampling can be treated with a restriction endonuclease (for example, Dde I and Mst II) that recognizes a sequence altered by the β^S mutation (8-11). This generates β^A - and β^S -specific restriction fragments that can be resolved by Southern transfer and hybridization with a β -globin probe.

We have developed a procedure for the detection of the sickle cell mutation that is very rapid and is at least two orders of magnitude more sensitive than standard Southern blotting. There are two special features to this protocol. The first is a method for amplifying specific β -globin DNA sequences with the use of oligonucleotide primers and DNA polymerase (12). The second is the analysis of the β -globin genotype by solution hybridization of the amplified DNA with a specific oligonucleotide probe and subsequent digestion with a restriction endonuclease (13). These two techniques increase the speed and sensitivity, and

lessen the complexity of prenatal diagnosis for sickle cell anemia; they may also be generally applicable to the diagnosis of other genetic diseases and in the use of DNA probes for infectious disease diagnosis.

Sequence amplification by polymerase chain reaction. We use a two-step procedure for determining the β -globin genotype of human genomic DNA samples. First, a small portion of the β -globin gene sequence spanning the polymorphic Dde I restriction site diagnostic of the β^A allele is amplified. Next, the presence or absence of the Dde I restriction site in the amplified DNA sample is determined by solution hybridization with an end-labeled complementary oligomer followed by restriction endonuclease digestion, electrophoresis, and autoradiography.

The β -globin gene segment was amplified by the polymerase chain reaction (PCR) procedure of Mullis and Faloona (12) in which we used two 20-base oligonucleotide primers that flank the region to be amplified. One primer, PC04, is complementary to the (+)-strand and the other, PC03, is complementary to the (-)-strand (Fig. 1). The annealing of PC04 to the (+)-strand of denatured genomic DNA followed by extension with the Klenow fragment of *Escherichia coli* DNA polymerase I and deoxynucleotide triphosphates results in the synthesis of a (-)-strand fragment containing the target sequence. At the same time, a similar reaction occurs with PC03, creating a new (+)-strand. Since these newly synthesized DNA strands are themselves template for the PCR primers, repeated cycles of denaturation, primer annealing, and extension result in the exponential accumulation of the 110-base pair region defined by the primers.

An example of the degree of specific gene amplification achieved by the PCR method is shown in Fig. 2A. Samples of DNA (1 μ g) were amplified for 20 cycles and a fraction of each sample, equivalent to 36 ng of the original DNA, was subjected to alkaline gel electrophoresis and transferred to a nylon filter. The filter was then hybridized with a 32 P-labeled 40-base oligonucleotide probe, RS06, which is complementary to the target sequence (Fig. 1A) but not to the PCR primers. The results, after a 2-hour autoradiographic exposure, show that a fragment hybridizing with the RS06 probe

The authors are in the Department of Human Genetics, Cetus Corporation, 1400 Fifty-Third Street, Emeryville, California 94608. The present address for N.A. is Department of Biological Sciences, University of Southern California, Los Angeles 90089-0371.

labeled fragment generated by cleavage of the probe.

For the diagnosis of sickle cell anemia, the probe was designed to be complementary to a region of the β -globin gene locus surrounding the sixth codon. In the β^A allele, the nucleotide (nt) sequence at this position contains a Dde I restriction site, but due to the single base mutation, this site is absent in the β^S allele. Our strategy for generating specific probe cleavage products for each allele is shown in Fig. 3. It is based on the presence of an invariant Hinf I restriction enzyme site immediately adjacent to the polymorphic Dde I restriction site. Resolution of the labeled oligomer cleavage products produced by sequential digestion with Dde I and Hinf I allows us to distinguish between the two alleles. In an individual homozygous for the wild-type β -globin allele AA, Dde I digestion will produce a labeled octamer (8 nt) from the probe. Because of its short length, the 8-nt cleavage product will dissociate from the genomic target DNA and the subsequent digestion with Hinf I has no effect. In the case of SS homozygotes, however, Dde I digestion does not cleave the probe since a base pair mismatch exists in the recognition sequence formed between the probe and target DNA. The invariant Hinf I site will then be cleaved during Hinf I digestion, releasing a labeled trimer (3 nt). In an AS heterozygote, both a trimer and an oc-

tamer would be detected. The resolution of the intact 40-base probe, the 8-nt and the 3-nt cleavage products is achieved by polyacrylamide gel electrophoresis. Experiments testing the sequential digestion strategy with plasmids carrying the β^A and β^S alleles show that, in each case, the expected probe cleavage products were produced (Fig. 4).

Analysis of genomic DNA samples by PCR and OR. Eleven DNA samples derived from lymphoblastoid cell lines or white blood cells were analyzed for their β -globin genotype by standard Southern blotting and hybridization of the Mst II RFLP (10), identifying the genotypes of the samples as either AA, AS, or SS. Six of these samples (and one additional one) were then amplified by PCR for 20 cycles starting with 1 μ g of DNA each. An aliquot of the amplified DNA sample (one-fourteenth of the original 1- μ g sample) was hybridized to the RS06 probe and digested with Dde I and then Hinf I. A portion (one-tenth) of this oligomer restriction reaction was analyzed on a polyacrylamide gel to resolve the cleavage products, and the results obtained after 6 hours of autoradiography are shown Fig. 5. The high sensitivity achieved with the PCR and OR method is demonstrated by the strength of the autoradiographic signal derived from only 1/140 of the original 1- μ g sample (7 ng). Each sample determined to be AA by RFLP analysis showed a strong 8-nt

fragment while those typed as SS showed a strong 3-nt fragment. Analysis of the known AS samples revealed both cleavage products.

In the analysis of the AA samples, a faint 3-nt could be detected in addition to the primary 8-nt signal. The reasons for this band remain unclear, although incomplete Dde I cleavage or the occasional failure of the 8-nt fragment to disassociate from the target DNA may contribute to the nonspecific 3-nt product generated by Hinf I digestion. In the analysis of the SS samples, a very faint 8-nt band was also observed in addition to the expected 3-nt signal. We have determined that the background 8-nt product detected in SS samples can be attributed to the δ -globin gene, which is highly homologous to β -globin. The nucleotide sequence of the two β -globin primers used for amplification is shown in Fig. 1. The downward pointing arrows indicate the differences between the β - and δ -globin genes. We hypothesized that the faint 8-nt signal observed in the SS samples was due to some amplification of the δ -globin gene by these primers and the subsequent cross-hybridization of the amplified δ sequences with the RS06 probe used in the OR procedure. δ -Globin has the same Dde I site as normal β -globin, and the duplex formed between an amplified δ gene segment and the RS06 probe would be expected to yield an 8-nt fragment on Dde I digestion even

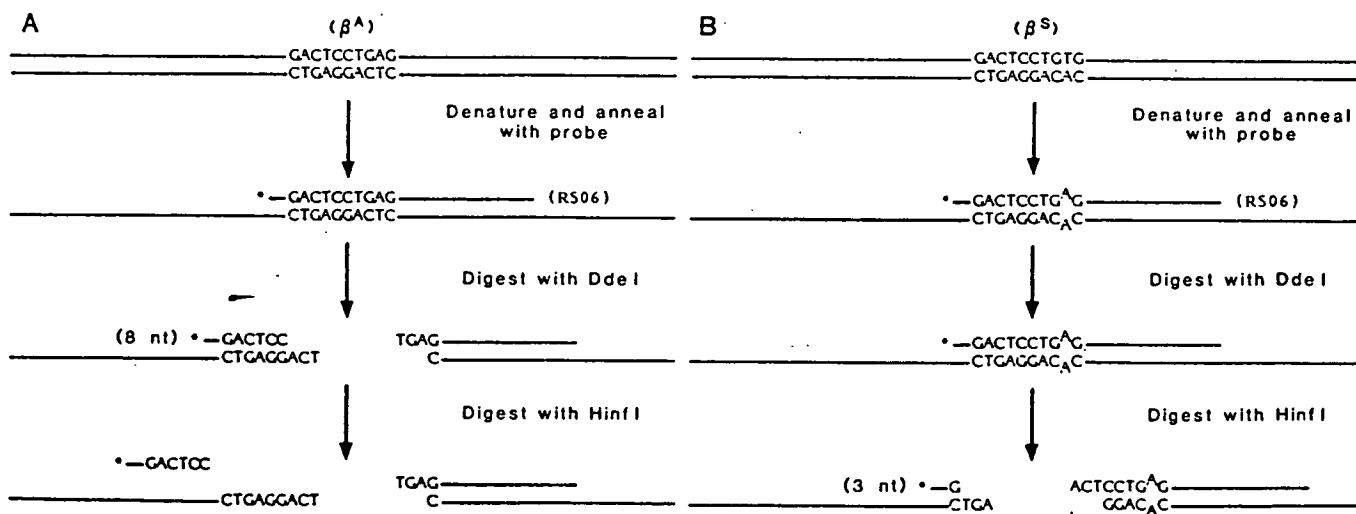


Fig. 3. Schematic diagram of oligomer restriction by sequential digestion to identify β^A - and β^S -specific cleavage products. The DNA sequences shown are the regions of the β -globin genomic DNA and the RS06 hybridization probe containing the invariant Hinf I site (GANTC, where N represents any nucleotide) and the polymorphic Dde I site (CTNAG). The remaining DNA sequences are represented as solid horizontal lines. The asterisk indicates the position of the radioactive 32 P label attached to the 5'-end of the RS06 probe with polynucleotide kinase. (A) Outline of the procedure and expected results when RS06 anneals to the normal β -globin gene (β^A). After denaturation of the genomic DNA and hybridization of the labeled RS06 probe to the complementary target sequence in the β^A gene, digestion of the probe-target hybrid with Dde I causes the release of a labeled (8-nt) cleavage product. Because of the relatively stringent conditions during Dde I digestion, the 8-nt cleavage product dissociates from the genomic DNA and the subsequent digestion with Hinf I has no effect. (B) Outline of Dde I and Hinf I digestion after hybridization of the RS06 probe to the sickle cell allele (β^S). As a consequence of the β^S mutation, the probe-target hybrid contains an A-A mismatch within the Dde I site and is not cleaved by the Dde I endonuclease. The Hinf I site, however, remains intact and digestion with that enzyme generates a labeled 3-nt product. Thus, the presence of the β^A allele is revealed by the release of a labeled 8-nt fragment, while the presence of β^S is indicated by a labeled 3-nt fragment.

though there are sequence differences (four mismatch out of 40 bases) between RS06 and δ -globin. It is likely that δ -globin sequences may be amplified to some extent and detected weakly with the RS06 probe in all DNA samples, but that its contribution to the total signal is very small and detectable only when the sample is SS and no 8-nt fragment from the β -globin gene is expected. We tested this hypothesis by treating an SS DNA sample before amplification with the enzyme Mbo I. Since there is a recognition site for this enzyme in the target DNA of the δ - but not the β -globin gene, cleavage of the δ gene between the regions that hybridize to the PCR primers should prevent its subsequent amplification (but not of β -globin). Our results showed that an SS DNA sample, first digested with Mbo I, gave only the 3-nt product but not the 8-nt product, this is consistent with the hypothesis of δ -globin amplification.

Effect of PCR cycle number on detection threshold. The strength of the autoradiograph signal detected by OR as a function of PCR cycle number and autoradiographic exposure was examined. The signal intensity after 20 cycles is at least 20 times as strong as that for 15 cycles and the determination of the β -globin genotype can be made with an autoradiographic exposure for only 2 hours (Fig. 6). The observed increase of ≥ 20 -fold is consistent with our estimates of 85 percent efficiency per cycle, calculated from the data in Fig. 2B ($1.85^5 = 21.7$). Coupled with the time that it takes to actually carry out the PCR and OR procedures, a 20-cycle PCR allows a diagnosis to be made in less than 10 hours with a DNA sample of 1 μ g.

Since all of the previous PCR experiments were done with 1 μ g of genomic DNA, we explored the effect of using significantly smaller amounts of DNA as template for PCR amplification. The results obtained with 20 cycles of PCR amplification on 500, 100, 20, and 4 ng of DNA from an AS individual are shown in Fig. 7. After analysis of 1/40 of each sample by the OR procedure and a 20-hour autoradiographic exposure, the β -globin genotype could be easily determined on DNA samples of 20 ng or about 100 times less than is needed for a typical Southern transfer and hybridization experiment. In this experiment, only a small fraction (1/40) of the starting material was placed on the gel; therefore it should be possible to analyze samples of less than 20 ng of genomic DNA (20 ng is equivalent to approximately 6000 haploid genomes) if a larger proportion of the material was utilized in the OR and gel electrophoresis steps.

Diagnostic applications of the PCR-OR system. When currently available methods are used, the completion of a prenatal diagnosis for sickle cell anemia takes a period of several days after the DNA is isolated. With 1 μ g of genomic DNA, the β -globin genotype can be determined by the PCR-OR method in less than 10 hours; 20 cycles of amplification requires about 2 hours (each full cycle takes 6 to 7 minutes in our protocol), the oligomer restriction procedure involving liquid hybridization and enzyme digestions require an additional 2 hours, and the electrophoresis takes about an hour. Autoradiographic exposure for 4 hours is sufficient to generate a strong signal.

Because this method includes a liquid hybridization protocol and involves the serial addition of reagents to a single tube, it is simpler to perform than the standard Southern transfer and hybridization procedure. Prior to electrophoresis, all of the reactions can be done in two small microcentrifuge tubes and could readily be automated.

The sensitivity, as well as the speed and simplicity, of this procedure is also important for clinical applications. Twenty nanograms of starting material can provide an easily detectable result in an overnight autoradiographic exposure. This sensitivity makes the PCR-OR method particularly valuable in cases

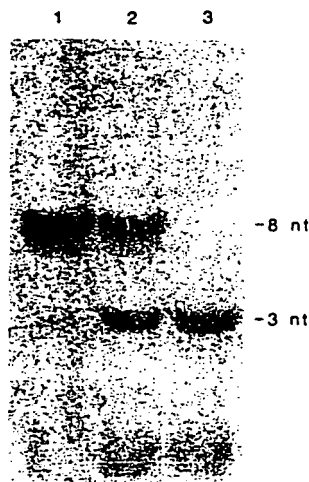


Fig. 4. Demonstration of OR sequential digestion with cloned β -globin genes. The sequential digestion strategy was demonstrated by annealing the RS06 probe to the β -globin plasmids pBR328:: β^A and pBR328:: β^S (13). The methods were similar to those described (13). Cloned β -globin DNA (45 ng; 0.01 pmol) was placed in a microcentrifuge tube, adjusted to 30 μ l with TE buffer (10 mM tris, 0.1 mM EDTA, pH 8.0), overlaid with 0.1 ml of mineral oil. The DNA was denatured by heating for 5 to 10 minutes at 95°C. Ten microliters of 0.6M NaCl containing 0.02 pmol of phosphorylated (with [γ - 32 P]ATP) RS06 probe oligomer (~5 μ Ci/pmol) was added and annealed for 60 minutes at 56°C. Unlabeled RS10 blocking oligomer (4 μ l; 200 pmol/ml) (Fig. 1) (13) was then added, and the hybridization was continued for 5 to 10 minutes. Next, 5 μ l of 100 mM MgCl₂ and 1 μ l of Dde I (Biolabs, 10 units) was added and incubated for 20 minutes at 56°C; 1 μ l of Hinf I (Biolabs, 10 units) was added and digestion was continued for 20 minutes at the same temperature. The reaction was terminated by the addition of 4 μ l of 100 mM EDTA and 6 μ l of tracking dye to a final volume of 61 μ l; a portion (8 μ l) (6 ng, 0.0013 pmol) was applied to a 0.75-mm thick, 30 percent polyacrylamide minigel (19 acrylamide:1 bis) in a Hoefer SE200 apparatus and subjected to electrophoresis (300 V) for 1 hour until the bromphenol blue dye front reached 3 cm. The top 1.5 cm of the gel, containing intact RS06, was cut off and discarded. The remaining gel was autoradiographed with a single intensification screen for 18 hours at -70°C. (Lane 1) six nanograms of pBR328:: β^A ; (lane 2) 3 ng of pBR28:: β^A plus 3 ng of pBR328:: β^S ; and (lane 3) 6 ng of pBR328:: β^S .

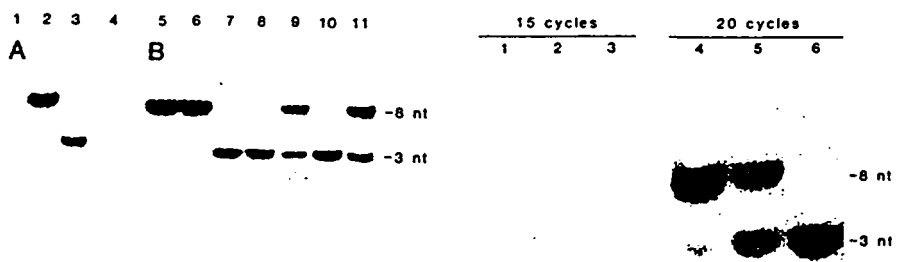


Fig. 5 (left). Determination of the β -globin genotype in human genomic DNA with PCR-OR. Samples (1 μ g) of human genomic DNA were amplified for 20 cycles (as described in Fig. 2A). The amplified DNA's (71 ng) were hybridized to the RS06 probe and serially digested with Dde I and Hinf I (as described in Fig. 4). Each sample (6 μ l) was analyzed by 30 percent polyacrylamide gel electrophoresis and autoradiographed for 6 hours at -70°C with one intensification screen. Each lane contains 7 ng of genomic DNA. (Lane 1) Unamplified Molt4 DNA (negative control); (lane 2) amplified Molt4 ($\beta^A\beta^A$); (lane 3) SC-1 ($\beta^S\beta^S$); (lane 4) GM2064 ($\Delta\Delta$); (lanes 5 to 11) clinical samples CH1 ($\beta^A\beta^A$), CH2 ($\beta^A\beta^A$), CH3 ($\beta^S\beta^S$), CH4 ($\beta^S\beta^S$), CH7 ($\beta^A\beta^A$), CH8 ($\beta^S\beta^S$), and CH12 ($\beta^A\beta^S$), respectively. Fig. 5 (right). Effect of cycle number on signal strength. Genomic DNA (1 μ g) from the clinical samples CH2 ($\beta^A\beta^A$), CH12 ($\beta^A\beta^S$), and CH5 ($\beta^S\beta^S$) were amplified for 15 and 20 cycles and equivalent amounts of genomic DNA (80 ng) were analyzed by oligomer restriction. (Lanes 1 to 3) DNA (20 ng) from CH2, CH12, and CH5, respectively, amplified for 15 cycles; (lanes 4 to 6) DNA (20 ng) from CH2, CH12, and CH5, respectively, amplified for 20 cycles. Autoradiographic exposure was for 2.5 hours at -70°C with one intensification screen.

where poor DNA yields are obtained from prenatal samples. In addition, DNA samples of poor quality (very low average molecular weight) can give excellent results in the PCR-OR protocol.

The PCR method is likely to be generally applicable for specific gene amplification since a fragment encoding a portion of the HLA-DQ α locus has recently been amplified with this procedure (15). We have carried out PCR amplification on a 110-bp β -globin sequence with an overall efficiency per cycle of about 85 percent. We have also amplified longer β -globin fragments (up to 267 bp), but the yield was lower under our standard conditions. Efficient amplification of a 267-bp fragment required some variation in the PCR procedure. In principle, increasing the number of PCR cycles should yield even greater amplification than that reported here (~220,000-fold after 20 cycles).

Our method for the diagnosis of sickle cell anemia involves the coupling of the PCR procedure with that of oligomer restriction. It was designed to distinguish between two alleles that differ by a polymorphic restriction site. The PCR-OR method is applicable as well to the diagnosis of other diseases where the lesion directly affects a restriction enzyme site or where the polymorphic site is in strong linkage disequilibrium with the disease causing locus. If the polymorphism is in linkage equilibrium with the disease, PCR-OR requires family studies to follow the inheritance of the disease locus.

In the case of the β -globin locus, the presence of the invariant Hinf I restriction site adjacent to the polymorphic Dde I site allows a sequential digestion procedure to identify both the β^A and β^S

Fig. 7. Detection threshold for PCR-OR. Fivefold serial dilutions of genomic DNA (500, 100, 20, and 4 ng) from the clinical sample CH12 ($\beta^A\beta^S$) were amplified by 20 cycles of PCR and one-tenth each reaction (50, 10, 2, and 0.4 ng) was analyzed by OR. The gel continued (lane 1) genomic DNA (12.5 ng); (lane 2) 2.5 ng; (lane 3) 0.5 ng; (lane 4) 0.1 ng; (lane 5) 12.5 ng genomic DNA from the globin deletion cell line GM2064. Autoradiographic exposure was for 20 hours at -70°C with an intensification screen.

alleles. In principle, this approach does not require that the two sites be immediately adjacent but only that the cleavage product generated by digestion at the polymorphic site dissociate from the target to prevent cutting at the invariant site. Since the restriction enzyme digestion conditions used here are fairly stringent for hybridization, we estimate that the polymorphic and invariant sites could perhaps be separated by as much as 20 bp.

The application of the PCR method to prenatal diagnosis does not necessarily depend on a polymorphic restriction site or on the use of radioactive probes. In fact, the significant amplification of target sequences achieved by the PCR method allows the use of nonisotopically labeled probes (16). Amplified target sequences could also be analyzed by a

number of other procedures including those involving the hybridization of small labeled oligomers which will form stable duplexes only if perfectly matched (6, 7, 17, 18) and the recently reported method based on the electrophoretic shifts of duplexes with base pair mismatches (19). The ability of the PCR procedure to amplify a target DNA segment in genomic DNA raises the possibility that its use may extend beyond that of prenatal diagnosis to other areas of molecular biology.

References and Notes

1. Y. W. Kan and A. M. Dozy, *Proc. Natl. Acad. Sci. U.S.A.* 75, 5631 (1978).
2. K. E. Davies and D. P. Ellis, *Biochem. J.* 226, 1 (1985).
3. Y. W. Kan, M. S. Golbus, A. M. Dozy, *N. Eng. J. Med.* 295, 1165 (1976).
4. A. M. Dozy et al., *J. Am. Med. Assoc.* 241, 1610 (1979).
5. E. M. Rubin and Y. W. Kan, *Lancet* 1985-I, 75 (1985).
6. M. Pirastu et al., *N. Engl. J. Med.* 309, 284 (1983).
7. S. H. Orkin, A. F. Markham, H. H. Kazazian, Jr., *J. Clin. Invest.* 71, 775 (1983).
8. R. F. Geever et al., *Proc. Natl. Acad. Sci. U.S.A.* 78, 5081 (1981).
9. J. C. Chang and Y. W. Kan, *N. Eng. J. Med.* 307, 30 (1982).
10. S. H. Orkin et al., *ibid.*, p. 32.
11. J. T. Wilson et al., *Proc. Natl. Acad. Sci. U.S.A.* 79, 3628 (1982).
12. F. Faloona and K. Mullis, in preparation.
13. R. Saiki, N. Arnheim, H. Erlich, *Biotechnology* 3, 1008 (1985).
14. D. Tuan et al., *Proc. Natl. Acad. Sci. U.S.A.* 80, 6937 (1983).
15. S. Scharf, unpublished data.
16. R. Saiki et al., unpublished data.
17. B. J. Conner et al., *Proc. Natl. Acad. Sci. U.S.A.* 80, 278 (1983).
18. V. J. Kidd, R. B. Wallace, K. Itakura, S. L. C. Woo, *Nature (London)* 304, 230 (1983).
19. R. M. Myers, N. Lumelsky, L. S. Lerman, T. Maniatis, *ibid.* 313, 495 (1985).
20. Calculation of the amounts of pBR328: β^A needed as standards to estimate PCR efficiency was done in the following way. If we assume that a human haploid genome size is 3×10^9 bp, 36 ng of DNA is equivalent to 1.8×10^{-8} pmol. The extent of amplification after 20 cycles at, for example, 85 percent efficiency is obtained with $(1.8 \times 10^{-8}) (1.85^{20})$ or 4.0×10^{-3} pmol of plasmid DNA (Fig. 2B, lane 5).

20 September 1985; accepted 15 November 1985